

Robust Distributed Estimation over Multiple Access Channels with Constant Modulus Signaling

Cihan Tepedelenlioğlu, *Member, IEEE*, Adarsh B. Narasimhamurthy

Abstract

A distributed estimation scheme where the sensors transmit with constant modulus signals over a multiple access channel is considered. The proposed estimator is shown to be strongly consistent for any sensing noise distribution in the i.i.d. case both for a per-sensor power constraint, and a total power constraint. When the distributions of the sensing noise are not identical, a bound on the variances is shown to establish strong consistency. The estimator is shown to be asymptotically normal with a variance (AsV) that depends on the characteristic function of the sensing noise. Optimization of the AsV is considered with respect to a transmission phase parameter for a variety of noise distributions exhibiting differing levels of impulsive behavior. The robustness of the estimator to impulsive sensing noise distributions such as those with positive excess kurtosis, or those that do not have finite moments is shown. The proposed estimator is favorably compared with the amplify and forward scheme under an impulsive noise scenario. The effect of fading is shown to not affect the consistency of the estimator, but to scale the asymptotic variance by a constant fading penalty depending on the fading statistics. Simulations corroborate our analytical results.

Index Terms

Distributed Estimation, Multiple Access Channel, Constant Modulus, Empirical Characteristic Function

The authors are with the School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ 85287, USA. (Email: {cihan,anbangal}@asu.edu).

I. INTRODUCTION

In inference-based wireless sensor networks, low-power sensors with limited battery and peak-power capabilities transmit their observations to a fusion center (FC) for detection of events or estimation of parameters. For distributed estimation, much of the literature has focused on a set of orthogonal (parallel) fading channels between the sensors and the FC (please see [1] and the references therein). The bandwidth requirements of such an orthogonal WSN scales linearly with the number of sensors. In contrast, over multiple access channels where the sensor transmissions are simultaneous and in the same frequency band, the utilized bandwidth does not depend on the number of sensors. In both cases, sensors may adopt either a digital or analog method for relaying the sensed information to the FC. The digital method consists of quantizing the sensed data and transmitting with digital modulation over a rate-constrained channel. In these cases, the required channel bandwidth is proportional number of bits at the output of the quantizer which are transmitted after pulse shaping and digital modulation. The analog method consists of transmitting unquantized data by appropriately pulse shaping and amplitude or phase modulating to consume finite bandwidth.

The literature on distributed estimation over multiple access channels has mainly involved analog sensor transmission schemes where the instantaneous transmit power is influenced by the sensor measurement noise and is not bounded [2]–[8]. In [2], distributed estimation over Gaussian multiple access channels is studied from a joint source-channel coding point of view. Reference [3] considers optimization of the sensor gains in the presence of channel fading. In [4] and [5], the effects of different fading distributions and channel feedback on the performance of distributed estimators over multiple access channels is studied. A direct-sequence CDMA with amplify and forward (AF) is considered in [6], where the asymptotic MSE is studied. In [7], the authors introduce a type-based multiple access scheme where more than one orthogonal channel is utilized albeit less in number than the number of sensors. In [8], a likelihood-based multiple access approach is introduced. The latter two references do not explicitly estimate a location parameter (such as the mean or the median) of the sensed data. In these aforementioned schemes, the sensor power management issues arising from the dependence of the instantaneous transmit power on the sensing noise have not been addressed. Moreover, for sensors operating

in adverse conditions, robustness to impulsive noise ¹ is of paramount importance, which has not been addressed in the literature in the context of distributed estimation over multiple access channels.

In this work, a distributed estimation scheme is considered where the sensor transmissions have constant modulus with fixed instantaneous transmit power. The proposed estimator is universal in the sense of [9] (or “distribution-free” in statistical parlance) in that the estimator does not depend on the distribution or the parameters of the sensing or channel noise. Unlike the orthogonal framework in [9], multiple access channels are considered herein, and the sensing noise is not assumed bounded. The estimator is shown to be strongly consistent for any noise distribution, including those with no finite moments, in the i.i.d. case. The distribution-free aspect is also very useful in heterogenous scenarios where several different kinds of noise are simultaneously present, such as additive Gaussian noise along with quantization noise.

The sensors transmit with constant modulus transmissions whose phase is linear with the sensed data. The FC estimates a common location parameter (such as the mean, or the median) of the sensed signal where the sensing noise samples are not assumed to be identically distributed, or from any specific distribution. It is shown that the proposed estimator is strongly consistent even when the sensing noise is not identically distributed, provided that their variances are bounded. While the estimator is shown to be consistent in this general framework, the asymptotic variance of the estimator is derived for the i.i.d. sensing noise and shown to depend on its characteristic function (CF). Upper bounds on, and optimization of the asymptotic variance with the transmit phase parameter ω is considered for different distributions on the sensing noise including impulsive ones. The proposed estimator is compared with AF, where the robustness of the proposed estimator is highlighted. The effect of fading is shown to not affect the consistency of the estimator, but only to scale the asymptotic variance by a constant fading penalty depending on the fading statistics.

II. SYSTEM MODEL

Consider the sensing model, with L sensors,

$$x_i = \theta + \eta_i \quad i = 1, \dots, L \quad (1)$$

¹referring to distributions whose tails decay slower than that of Gaussian noise

where θ is an unknown real-valued parameter in a bounded interval $[0, \theta_R]$ of known length, $\theta_R < \infty$, η_i are a mutually independent, symmetric real-valued noise with zero median (i.e., its pdf, when it exists, is symmetric about zero), and x_i is the measurement at the i^{th} sensor. Note that η_i are not necessarily identically distributed, bounded, and need not have finite moments. We consider a setting where the i^{th} sensor transmits its measurement using a constant modulus signal $\sqrt{\rho}e^{j\omega x_i}$ over a Gaussian multiple access channel so that the received signal at the fusion center (FC) is given by

$$y_L = \sqrt{\rho} \sum_{i=1}^L e^{j\omega x_i} + v \quad (2)$$

where the transmitted signal at each sensor has a per-sensor power of ρ , $0 < \omega \leq 2\pi/\theta_R$ is a design parameter to be optimized, and v is additive noise. Note that the restriction $\omega \in (0, 2\pi/\theta_R]$ is necessary even in the absence of sensing and channel noise, to uniquely determine θ from y_L . Estimation in a single time snap shot is considered, which is why the time index is dropped. The transmitted signal has a deterministic fixed power ρ which does not suffer from the problems of random transmit power seen in AF schemes where the transmitted signal from the i^{th} sensor is given by $\alpha x_i = \alpha(\theta + \eta_i)$ with instantaneous power per sensor $\alpha^2(\theta + \eta_i)^2$, which is an unbounded random variable (RV) when η_i is. In AF transmission, α is a coefficient which might depend on the sensor index, as well as on L through a power constraint, but does not depend on x_i [10], [11]. Note that the total transmit power from all the sensors in (2) is ρL . We begin by considering a fixed total power constraint P_T implying that the per-sensor power $\rho = P_T/L$ is a function of L . Later, in Section IV-A, we will also consider a fixed per-sensor power scheme where ρ will not be a function of the number of sensors L , which implies $P_T \rightarrow \infty$ as $L \rightarrow \infty$.

III. THE ESTIMATION PROBLEM

We would like to estimate θ from y_L which under the total power constraint is given by

$$y_L = e^{j\omega\theta} \sqrt{\frac{P_T}{L}} \sum_{i=1}^L e^{j\omega\eta_i} + v. \quad (3)$$

We do not assume that η_i are identically distributed, or that η_i are from any specific distribution since a universal estimator which is independent of the distribution of η_i is desired. Let,

$$z_L := \frac{y_L}{\sqrt{L}} = e^{j\omega\theta} \sqrt{P_T} \frac{1}{L} \sum_{i=1}^L e^{j\omega\eta_i} + \frac{v}{\sqrt{L}}, \quad (4)$$

and define $\varphi_{\eta_i}(\omega) := \mathbb{E}[e^{j\eta_i\omega}]$ as the CF of η_i . Due to the law of large numbers we have

$$\frac{1}{L} \sum_{i=1}^L e^{j\eta_i\omega} \rightarrow \varphi(\omega) := \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{i=1}^L \varphi_{\eta_i}(\omega) \quad (5)$$

(where \rightarrow indicates convergence almost surely), and we use the fact that the variances $\text{var}(e^{j\eta_i\omega}) = 1 - \varphi_{\eta_i}^2(\omega) \leq 1$ are bounded to invoke Kolmogorov's strong law of large numbers for non-identically distributed RVs [12, pp. 259]. Since η_i are symmetric, $\{\varphi_{\eta_i}(\omega)\}$ are real-valued and therefore $\varphi(\omega)$ is also real-valued.

Consider the conditions under which $\varphi(\omega)$ is a CF, which will be important in the consistency of the proposed estimator. Since convex combinations of CFs are CFs [13], the partial sums $L^{-1} \sum_{i=1}^L \varphi_{\eta_i}(\omega)$ are as well. From the continuity theorem [13, Corollary 1.2.2] if a sequence of CFs converges pointwise to a function continuous at $\omega = 0$, then the limit is a CF. Therefore $\varphi(\omega)$ in (5) is a CF if $\varphi(\omega)$ is continuous at $\omega = 0$.

The natural estimator that we will adopt is based on the phase of z_L :

$$\hat{\theta} = \frac{1}{\omega} \tan^{-1} \left(\frac{z_L^I}{z_L^R} \right), \quad (6)$$

where $z_L^R := \text{Re}\{z_L\}$ and $z_L^I := \text{Im}\{z_L\}$. Note that this estimator does not depend on the distributions of η_i or v , as desired. We now establish the strong consistency of the proposed estimator $\hat{\theta}$:

Theorem 1. *The estimator $\hat{\theta}$ in (6) is strongly consistent provided that $\omega \in (0, 2\pi/\theta_R]$ is chosen to satisfy $\varphi(\omega) \neq 0$.*

Proof: Taking the real and imaginary parts of (4) and (5) due to the strong law of large numbers $z_L^R \rightarrow \bar{z}^R := \sqrt{P_T} \cos(\omega\theta) \varphi(\omega)$ and $z_L^I \rightarrow \bar{z}^I := \sqrt{P_T} \sin(\omega\theta) \varphi(\omega)$ almost surely. Since $\hat{\theta}$ in (6) is a continuous function of $[z_L^R, z_L^I]$, $\hat{\theta} \rightarrow (1/\omega) \tan^{-1}(\bar{z}^I/\bar{z}^R) = \theta$ almost surely [14, Thm 3.14]. We need the assumption that $\varphi(\omega) \neq 0$ since otherwise θ cannot be uniquely determined from \bar{z}^R and \bar{z}^I . ■

We now investigate when an ω that satisfies the conditions of Theorem 1 exists. Consider first the identically distributed case where η_i have a common distribution with a RV η so that

$\varphi(\omega) = \varphi_\eta(\omega)$ is a CF. Many distributions such as Gaussian, Laplace, and Cauchy satisfy $\varphi_\eta(\omega) > 0$ for all ω . If the common sensing noise distribution is known to have this property, then any choice of $\omega \in (0, 2\pi/\theta_R]$ would clearly satisfy the conditions of Theorem 1. In the more general case, where nothing is known or assumed about η , a sufficiently small ω satisfies $\varphi(\omega) > 0$ since all CFs at the origin are equal to 1 and continuous. So, for identically distributed sensing noise, an ω for which (6) is strongly consistent can always be found, even if the sensing noise variance does not exist.

In the general non-identically distributed case, this argument does not follow since $\varphi(\omega)$ is not necessarily a CF. However, if $\varphi(\omega)$ is continuous at $\omega = 0$, it is a CF by the continuity theorem [13] and the argument above follows. For an example of when $\varphi(\omega)$ is not a CF and not continuous at $\omega = 0$, consider a case where $\sum_{i=1}^{\infty} \varphi_{\eta_i}(\omega) < \infty$ for all $\omega > 0$ such as when η_i are Gaussian with variances that depend on i linearly: $\varphi_{\eta_i}(\omega) = e^{-\sigma_i^2 \omega^2 / 2}$ where $\sigma_i^2 = i\sigma^2$, and $\sum_{i=1}^{\infty} \varphi_{\eta_i}(\omega) = (1 - \exp(-\sigma^2 \omega^2 / 2))^{-1} < \infty$ by the geometric sum formula. In this case due to the L^{-1} factor in (5), $\varphi(\omega) = 0$ when $\omega > 0$, and $\varphi(0) = 1$. For this example, $\varphi(\omega)$ is not a CF for any distribution, and there exists no ω that satisfies the requirements of Theorem 1. Clearly, this is a very severe case where the sensing noise variance increases linearly with the sensor index, without bound. In fact, the example above can be generalized to distributions other than Gaussian, and variances going to infinity even slower than linearly. For absolutely continuous sensing noise distributions, when η_i are expressed as a scalar multiple of an underlying random variable, and these scalars (which are proportional to standard deviations when they exist) go to infinity, it can be shown that the estimator in (6) is not consistent, which is proved next.

Theorem 2. *Let the sensing noise at the i^{th} sensor be a scaled version of a RV η with absolutely continuous distribution so that $\eta_i = \sigma_i \eta$ and $\varphi_{\eta_i}(\omega) = \varphi_\eta(\sigma_i \omega)$. Suppose also that $\lim_{i \rightarrow \infty} \sigma_i = \infty$. Then there is no ω that satisfies the conditions of Theorem 1.*

Proof: Recalling from (5) the definition of $\varphi(\omega)$, we would like to show that $\varphi(\omega) := \lim_{L \rightarrow \infty} L^{-1} \sum_{i=1}^L \varphi_\eta(\sigma_i \omega) = 0$ for $\omega > 0$. Since η has an absolutely continuous distribution, $\lim_{x \rightarrow \infty} \varphi_\eta(x) = 0$, and because $\lim_{i \rightarrow \infty} \sigma_i = \infty$, it follows that $\lim_{i \rightarrow \infty} \varphi_\eta(\sigma_i \omega) = 0$ for $\omega > 0$. From [14, pp. 411] we know that if a sequence satisfies $\lim_{i \rightarrow \infty} a_i = 0$ then $\lim_{L \rightarrow \infty} L^{-1} \sum_{i=1}^L a_i = 0$, which gives us the proof when applied to the sequence $\varphi_\eta(\sigma_i \omega)$. ■

The following theorem can loosely be regarded as a converse to Theorem 2 and shows that the estimator in (6) is consistent when the variances σ_i^2 exist and are bounded.²

Theorem 3. *Let $\text{var}(\eta_i)$ exist for all i and $\sigma_{\max} := \sup_i (\text{var}(\eta_i))^{1/2}$ be finite. Then any $0 < \omega < \min(2\pi/\theta_R, \sqrt{2}/\sigma_{\max})$ satisfies $\varphi(\omega) > 0$, thereby fulfilling the requirement of Theorem 1 on ω .*

Proof: From [13, pp. 89] we have $\varphi_{\eta_i}(\omega) \geq 1 - \sigma_i^2 \omega^2 / 2$ for any CF with finite variance. Using (5) we have $\varphi(\omega) \geq 1 - (\lim_{L \rightarrow \infty} L^{-1} \sum_{i=1}^L \sigma_i^2) \omega^2 / 2 \geq 1 - \sigma_{\max}^2 \omega^2 / 2 > 0$ where the last inequality holds provided that $\omega < \sqrt{2}/\sigma_{\max}$. Since also $\omega \leq 2\pi/\theta_R$ we have the theorem. ■

The estimator in (6) relies on constant modulus transmissions from the sensors to the FC, and is strongly consistent over a wide range of scenarios outlined above. However, the performance of $\hat{\theta}$ will depend on statistical assumptions on $\{\eta_i\}$ and v . The following theorem characterizes this performance, under the assumption that $v \sim \mathcal{CN}(0, \sigma_v^2)$ and $\{\eta_i\}$ are identically distributed with an arbitrary common distribution.

Theorem 4. $\sqrt{L}(\hat{\theta} - \theta)$ is asymptotically normal with zero mean and variance given by,

$$AsV(\omega) = \frac{\left[\frac{\sigma_v^2}{P_T} + 1 - \varphi_{\eta}(2\omega) \right]}{2\omega^2 \varphi_{\eta}^2(\omega)} \quad (7)$$

Proof: Please see Appendix 1. ■

Note that in the i.i.d. case (4) is the empirical characteristic function (ECF) [13] of η_i corrupted by additive noise. While the ECF has been studied extensively in the statistical literature for constructing centralized estimators [13], it has not been addressed in the context of communication of samples as in distributed estimation, and therefore issues of power constraint and channel noise have not arisen in the literature on parameter estimation with ECFs.

IV. ANALYSIS AND OPTIMIZATION OF THE AsV

The proposed estimator is consistent under general conditions and does not depend on the noise parameters. However, if the noise distribution and parameters are available, it is possible

²It is not a true converse for two main reasons: (i) $\lim_{i \rightarrow \infty} \sigma_i = \infty$ required by Theorem 2 is not the opposite of σ_i being bounded, which is required by Theorem 3, since it is possible that neither may occur; (ii) Theorem 2 requires absolute continuity whereas Theorem 3 does not.

to minimize the AsV with respect to ω over the interval $(0, 2\pi/\theta_R]$:

$$AsV^* := \inf_{\omega \in (0, 2\pi/\theta_R]} \frac{\left[\frac{\sigma_v^2}{P_T} + 1 - \varphi_\eta(2\omega) \right]}{2\omega^2 \varphi_\eta^2(\omega)}. \quad (8)$$

We will consider this problem with both per-sensor, and total power constraints.

A. Per-sensor Power Constraint

Our derivation for the estimator $\hat{\theta}$ in (6), its strong consistency in Theorem 1, and the asymptotic variance in (7) had assumed that P_T is fixed as a function of L . In the fixed per-sensor power constraint case the total power $P_T = \rho L$ increases linearly with L in which case the estimator is given in (6) with $z_L := y_L/L$ which we redefine with an extra factor of $1/\sqrt{L}$ in (4). In this case, the statement of Theorem 1 still holds exactly, with minor modifications in the proof, and $\sigma_v^2/P_T \rightarrow 0$ as $L \rightarrow \infty$. Hence, having a per-sensor power constraint is asymptotically equivalent to having no channel noise. In either case (8) becomes,

$$AsV_{\text{pspc}}^* = \inf_{\omega \in (0, 2\pi/\theta_R]} \frac{[1 - \varphi_\eta(2\omega)]}{2\omega^2 \varphi_\eta^2(\omega)}, \quad (9)$$

which is a special case of (8). The reason we consider this case separately is because, as we will see, the objective in (9) is bounded near the origin which makes the solution of (9) considerably different than that of (8). We now consider solving (9), and investigate the behavior of $AsV(\omega)$ near the origin to see under what conditions small ω will yield optimum performance. Using l'Hôpital's rule, it is seen that $\lim_{\omega \rightarrow 0} AsV(\omega) = \sigma_\eta^2$ the variance of η , when η has finite variance. In fact, when also the fourth moment μ_4 of η exists, we have a stronger result:

Theorem 5. *If the first four moments of η exists, then $AsV(\omega)$ in equation (9) satisfies*

$$AsV(\omega) = \sigma_\eta^2 - \frac{1}{3} \kappa_\eta \sigma_\eta^4 \omega^2 + o(\omega^2) \quad (10)$$

as $\omega \rightarrow 0$, where $\kappa_\eta := \mu_4/\sigma_\eta^4 - 3$ is the excess kurtosis of η .

Proof: We have already established that the first term in (10) is σ_η^2 . Using the Maclaurin series expansion of $\varphi_\eta(\omega)$ in terms of the second and fourth moments of η , the numerator and denominator of (9) can be expressed as $N(\omega) := 2\sigma_\eta^2\omega^4 + (2/3)\mu_4\omega^4 + o(\omega^4)$ and $D(\omega) := 2\omega^2 (1 - (1/2)\sigma_\eta^2\omega^2 + (\mu_4/24)\omega^4 + o(\omega^4))$, respectively. By taking the second derivative and evaluating we have

$$\left. \frac{\partial^2}{\partial \omega^2} \frac{N(\omega)}{D(\omega)} \right|_{\omega=0} = -\frac{2}{3} \mu_4 + 2\sigma_\eta^4. \quad (11)$$

Dividing by $2!$ we obtain the coefficient of ω^2 in the Maclaurin series, as given in (10). ■

Theorem 5 has some interesting implications. By making ω sufficiently small, we can obtain an AsV that is arbitrarily close to σ_η^2 . Also, if the excess kurtosis of the sensing noise is positive, it is possible to improve the AsV to a value smaller than σ_η^2 by increasing ω in the neighborhood of 0, which shows that if $\kappa_\eta > 0$, (9) satisfies $AsV_{\text{pspc}}^* < \sigma_\eta^2$. This is the case for impulsive distributions like the Laplace distribution where $\kappa_\eta = 3$. When η is Gaussian, the excess kurtosis $\kappa_\eta = 0$ and therefore it is not clear from (10) if $AsV_{\text{pspc}}^* < \sigma_\eta^2$ is possible, since (10) only applies near $\omega = 0$. The following theorem sheds more light on this issue.

Theorem 6. *If η is Gaussian then the best asymptotic performance for $\hat{\theta}$ in (6) for the per-sensor power constraint satisfies $AsV_{\text{pspc}}^* = \sigma_\eta^2$.*

Proof: Equation (10) shows that $\lim_{\omega \rightarrow 0} AsV(\omega) = \sigma_\eta^2$ which implies that $AsV_{\text{pspc}}^* \leq \sigma_\eta^2$. To see that $AsV_{\text{pspc}}^* \geq \sigma_\eta^2$ consider a benchmark genie-aided sample mean estimator $\hat{\theta}_{GA} = L^{-1} \sum_{i=1}^L x_i$ that has access to the sensor measurements $\{x_i\}_{i=1}^L$, rather than the normalized channel output z_L in (4). The sample mean which has an asymptotic variance of σ_η^2 achieves the Cramer Rao bound (CRB) for an estimator of θ from $\{x_i\}_{i=1}^L$ since it is an efficient estimator of the mean when η is Gaussian. Since $\theta \rightarrow \{x_i\}_{i=1}^L \rightarrow z_L$ forms a Markov chain, from the data processing inequality for the CRB [15], the CRB for estimators of θ based on z_L is at least that obtained for the genie-aided setup of estimating θ from $\{x_i\}_{i=1}^L$, which is σ_η^2 . Therefore, the best achievable performance in the per-sensor power case cannot be better than that of $\hat{\theta}_{GA}$, which implies $AsV_{\text{pspc}}^* \geq \sigma_\eta^2$. ■

Note that in the proof of Theorem 6 we used the Gaussianity only to assert that the sample mean achieves the CRB. Theorem 6 also holds for any other distribution with this property.

In Figure 1 we show the analytical expressions for $AsV(\omega)$ for various distributions. For the Cauchy distribution the performance is unbounded at the origin since the variance does not exist. For all other distributions, we selected $\sigma_\eta^2 = 1$, which is the value of $AsV(\omega)$ near the origin. Note that the Laplace distribution which has a positive excess kurtosis corresponds to an AsV which is decreasing near the origin, as predicted by (10), whereas the Gaussian and uniform distributions are increasing near the origin from their infimum value of $\sigma_\eta^2 = 1$.

To conclude, for this per-sensor power constraint case, small ω yields good asymptotic performance which does not depend on θ_R . The performance can be improved by appropriately

increasing ω in the neighborhood of $\omega = 0$ when η is from an impulsive distribution with positive excess kurtosis.

B. Total Power Constraint

In this case P_T is not a linear function of L as it was in the previous section, but a constant so that the AsV is given by (7). Note that small ω should be avoided in the solution of (8) since $\lim_{\omega \rightarrow 0} AsV(\omega) = \infty$ is no longer finite, as seen also in Figure 2 for various fading distributions. For the same reason, one may use min instead of sup in (8) for the total power constraint case, since the minimum is always achieved by a strictly positive ω , when $\sigma_v^2 > 0$.

1) *Upper bound on AsV :* In what follows, we use the lower bound $\varphi_\eta(\omega) \geq 1 - \sigma_\eta^2 \omega^2 / 2$ in order to upper bound AsV^* in (8). We have the following theorem which applies when θ_R is large enough so that $(2\pi\sigma_\eta)/\theta_R < \sqrt{2}$.

Theorem 7. *The best achievable performance AsV^* in (8) for any sensing noise distribution with finite variance σ_η^2 satisfies*

$$AsV^* \leq \frac{\left[\frac{4\sigma_v^2}{P_T} + c \right]}{c \left[1 - \frac{1}{16}c \right]^2} \quad (12)$$

whenever $c/8 < (2\pi/\theta_R)^2 \sigma_\eta^2 < 2$, where $c := -3\sigma_v^2/P_T + (\sigma_v/\sqrt{P_T})\sqrt{32 + 9\sigma_v^2/P_T}$. On the other hand, when $(2\pi/\theta_R)^2 \sigma_\eta^2 < c/8$ then,

$$AsV^* \leq \frac{\left[\frac{\sigma_v^2}{P_T} + 2 \left(\frac{2\pi}{\theta_R} \sigma_\eta \right)^2 \right]}{2 \left(\frac{2\pi}{\theta_R} \right)^2 \left(1 - \left(\frac{2\pi}{\theta_R} \sigma_\eta \right)^2 \frac{1}{2} \right)^2} \quad (13)$$

Proof: Please see Appendix 2. ■

Note that if the range of the unknown parameter determined by θ_R is large, the upper bound in (13) will be tighter since the bound $\varphi_\eta(\omega) \geq 1 - \sigma_\eta^2 \omega^2 / 2$ is tighter when ω is small. Moreover, when θ_R is large, (13) simplifies to $(\sigma_v^2/P_T)(\theta_R^2/8\pi^2) + \sigma_\eta^2$. This shows that if the range θ_R increases, the optimal achievable performance AsV^* increases as well. In addition to large θ_R , when $\sigma_v^2/P_T \rightarrow 0$ as in the per-sensor power constraint case, the bound further simplifies to $AsV^* \leq \sigma_\eta^2$. The bound in Theorem 7 holds regardless of the distribution on η as long as it has finite variance.

Instead of working with bounds, if exact solutions to (8) are desired, then it is necessary to specify the sensing noise distribution. In what follows, the problem is specialized by considering some common distributions. The resulting asymptotic variances for the different distributions are illustrated in Figure 2.

2) *Gaussian Sensing Noise:* In this case, we have $\varphi_\eta(\omega) = \exp(-\sigma_\eta^2 \omega^2 / 2)$ so that

$$AsV_G(\omega) = \frac{e^{\sigma_\eta^2 \omega^2}}{2\omega^2} \left[\frac{\sigma_v^2}{P_T} + 1 - e^{-2\sigma_\eta^2 \omega^2} \right]. \quad (14)$$

We would like to minimize (14) over $\omega \in (0, 2\pi/\theta_R]$ as in (8). As an intermediary step, we first characterize the unconstrained minimum over $\omega \in [0, \infty)$. To simplify (14) we substitute $\beta \leftarrow \sigma_\eta^2 \omega^2$. Note that the value of ω that minimizes (14) over $\omega > 0$ is related to the $\beta > 0$ that minimizes $AsV_G(\sqrt{\beta}/\sigma_\eta)$ through $\omega = \sqrt{\beta}/\sigma_\eta$. Differentiating with respect to β , we have,

$$\frac{\partial AsV_G(\sqrt{\beta}/\sigma_\eta)}{\partial \beta} = \frac{e^{-\beta} \left[\left(\frac{\sigma_v^2}{P_T} + 1 \right) (\beta - 1) e^{2\beta} + \beta + 1 \right] \sigma_\eta^2}{2\beta^2}. \quad (15)$$

Any stationary point of $AsV_G(\sqrt{\beta}/\sigma_\eta)$, with respect to β satisfies,

$$\left(\frac{\sigma_v^2}{P_T} + 1 \right) (\beta - 1) e^{2\beta} + (\beta + 1) = 0. \quad (16)$$

Let any solution to (16) be denoted as β_G^* . It is straightforward to show that $\partial^2 AsV_G(\sqrt{\beta}/\sigma_\eta) / \partial \beta^2|_{\beta=\beta_G^*}$ is positive. This proves that β_G^* is the unique unconstrained minimum of $AsV_G(\sqrt{\beta}/\sigma_\eta)$ over $\beta > 0$ which in turn implies that $\omega_G^* = \sqrt{\beta_G^*}/\sigma_\eta$ is the corresponding unique minimizer of $AsV_G(\omega)$ for $\omega > 0$. Since $AsV_G(\omega)$ has a unique minimum, it is monotonically decreasing over $\omega \in (0, \sqrt{\beta_G^*}/\sigma_\eta]$. The solution to (8) in the Gaussian case therefore is

$$\omega_G^* = \min \left(\frac{2\pi}{\theta_R}, \frac{\sqrt{\beta_G^*}}{\sigma_\eta} \right), \quad (17)$$

where β_G^* is the unique solution to (16).

While there has been some efforts in the physics community [16] to define functions that solve the intersection point of rational functions and exponentials as in (16), there is no widely accepted formula. But (16) can be easily solved numerically to optimize ω when η is Gaussian.

3) *Cauchy Sensing Noise:* For the Cauchy distribution $\varphi_\eta(\omega) = e^{-\gamma\omega}$ for $\omega > 0$. It is well known that no moments of this distribution exists. Substituting $\varphi_\eta(\omega)$ in (7), we have

$$AsV_C(\omega) = \frac{e^{2\gamma\omega}}{2\omega^2} \left[\frac{\sigma_v^2}{P_T} + 1 - e^{-2\gamma\omega} \right]. \quad (18)$$

As in the Gaussian case, we first find the stationary points of (18) on $\omega \in [0, \infty)$ by taking the derivative of (18) and equating to zero to obtain,

$$\beta_C^* = \frac{1}{2\gamma} \left[2 + W \left(\frac{-2P_T}{\sigma_v^2 + P_T} e^{-2} \right) \right] \quad (19)$$

where $W(\cdot)$ is the Lambert function defined to be the inverse function of xe^x . It can be verified that $AsV''(\beta_C^*) > 0$ and therefore β_C^* is the unique unconstrained minimum of $AsV_C(\omega)$. Hence, $AsV_C(\omega)$ has a unique minimum over $\omega > 0$, and the solution to (8) in this case is

$$\omega_C^* = \min \left(\frac{2\pi}{\theta_R}, \beta_C^* \right). \quad (20)$$

4) *Laplace Sensing Noise:* In this case, we have $\varphi_\eta(\omega) = (1 + b^2\omega^2)^{-1}$ where $b^2 := \sigma_\eta^2/2$. Substituting $\beta \leftarrow b^2\omega^2$ for convenience, (7) for Laplace noise becomes,

$$AsV_L \left(\sqrt{\frac{\beta}{b^2}} \right) = b^2 \frac{(1 + \beta)^2}{2\beta} \left[\frac{\sigma_v^2}{P_T} + \frac{4\beta}{1 + 4\beta} \right]. \quad (21)$$

To characterize the stationary points of (21), we take the derivative with respect to β and equate to zero. The optimum value is the root of a 4th order polynomial. Using the only solution with a positive root we have,

$$\beta_L^* = \frac{1}{12} \left(\frac{c}{\frac{\sigma_v^2}{P_T} + 1} + \frac{25\frac{\sigma_v^2}{P_T} + 4}{c} + 2 \right) \quad (22)$$

where

$$c = \left[125 \left(\frac{\sigma_v^2}{P_T} \right)^3 + 258 \left(\frac{\sigma_v^2}{P_T} \right)^2 + 141 \left(\frac{\sigma_v^2}{P_T} \right) + 3\sqrt{3} \sqrt{\left(\frac{\sigma_v^2}{P_T} \right) \left(\frac{\sigma_v^2}{P_T} + 1 \right)^3 \left(375 \frac{\sigma_v^2}{P_T} + 32 \right)} + 8 \right]^{1/3}.$$

It is also possible to verify that the second derivative is positive at the optimal point. To express the roots of the 4th order polynomial in closed-form and verifying that the second derivative is always positive, we have used Mathematica. Using (22), $\omega^2 b^2 = \beta$, and the fact that $\omega \in (0, 2\pi/\theta_R]$ we have the solution to (8) as

$$\omega_L^* = \min \left(\frac{2\pi}{\theta_R}, \frac{1}{b} \sqrt{\beta_L^*} \right). \quad (23)$$

5) *Uniform Sensing Noise:* We now assume that η is uniformly distributed on $[-a, a]$, so that $a^2 = 3\sigma_\eta^2$. In this case $\varphi_\eta(\omega) = \sin(\omega a)/(\omega a)$ and we need to optimize,

$$AsV_U(\omega) = \frac{a^2}{2 \sin^2(\omega a)} \left[\frac{\sigma_v^2}{P_T} + \left(1 - \frac{\sin(2\omega a)}{2\omega a} \right) \right] \quad (24)$$

over $\omega \in (0, 2\pi/\theta_R]$. Note that $AsV_U(\omega)$ is undefined at $\omega = \pi/a$. We begin by showing that the range of ω can be further reduced to $\omega \in (0, \min(2\pi/\theta_R, \pi/a))$ in solving (8). This is because both $\sin^2(\omega a)$ and $\sin(2\omega a)$ are periodic with period π/a , and therefore due to the $2\omega a$ term in the denominator, $AsV_U(\omega) \leq AsV_U(\omega + k\pi/a)$ for any positive integer k , and $\omega > 0$.

In order to minimize (24) over $\omega \in (0, \min(2\pi/\theta_R, \pi/a))$, we first disregard the constraint on ω imposed by θ_R , and focus on $\omega \in (0, \pi/a)$. Substituting $\beta \leftarrow \omega a$, differentiating $AsV_U(\beta/a)$ with respect to β and equating to zero we obtain

$$\left[8 \left(\frac{\sigma_v^2}{P_T} + 1 \right) \beta^2 - 1 \right] \cos(\beta) + \cos(3\beta) - 4\beta \sin(\beta) = 0. \quad (25)$$

By taking the second derivative, it can be verified that $AsV_U(\beta/a)$ is convex, and therefore (25) has a unique solution β_U^* over $\beta \in (0, \pi)$ corresponding to the unique minimum of $AsV_U(\beta/a)$ over the same interval. It is immediate that $\omega = \beta_U^*/a$ is the unique minimum of (24) over $\omega \in (0, \pi/a)$, and therefore (24) is a monotonically decreasing function over $(0, \beta_U^*/a)$. Incorporating the effect of θ_R , we have that if $2\pi/\theta_R \leq \beta_U^*/a$ then the minimum of (24) over $\omega \in (0, \min(2\pi/\theta_R, \pi/a))$ is attained at $\omega = 2\pi/\theta_R$, and if $2\pi/\theta_R \geq \beta_U^*/a$, then it is attained at $\omega = \beta_U^*/a$. In short,

$$\omega_U^* = \min \left(\frac{2\pi}{\theta_R}, \frac{\beta_U^*}{a} \right) \quad (26)$$

Note that a closed-form solution to (25) is not possible, however a numerical solution can be easily found. Recall also from Section IV that for uniform noise which has $\kappa_\eta = -6/5$, a small $\omega > 0$ should be chosen when $\sigma_v^2/P_T = 0$. If instead $\sigma_v^2/P_T > 0$, then $\omega \approx \pi/2a$ (or $2\pi/\theta_R$, whichever one is smaller) is a good choice. We will elaborate on this more in Section IV-C, where we consider the low channel SNR regime.

6) *Compound Gaussian Sensing Noise:* Compound Gaussian is a class of RVs which when conditioned on the variance is a Gaussian RV. So when η is compound Gaussian, it can be written as $\eta = \sqrt{X}G$ where G is a Gaussian RV with zero mean and variance one, and X is a positive RV. It is easy to show that the CF of η can be expressed in terms of the moment

generating function (MGF) of X :

$$\varphi_\eta(\omega) = \mathbb{E} \left[e^{-\frac{1}{2}X\omega^2} \right] = M_X \left(-\frac{1}{2}\omega^2 \right) \quad (27)$$

where $M_X(t) := \mathbb{E}[e^{tX}]$ is the MGF of X when the expectation exists. Note that $\mathbb{E}[X] = \sigma_\eta^2$ in general and if the CDF of X is a unit step at σ_η^2 then η is Gaussian with variance σ_η^2 . For compound Gaussian sensing noise, (27) can be substituted in (7) to obtain

$$AsV_{CG}(\omega) = \frac{\left[\frac{\sigma_v^2}{P_T} + 1 - M_X(-2\omega^2) \right]}{2\omega^2 M_X^2(-\frac{1}{2}\omega^2)} \quad (28)$$

whenever the MGF exists.

When the per-sensor power is fixed so that $\sigma_v^2/P_T \rightarrow 0$ as $L \rightarrow \infty$, (28) can be expanded near $\omega = 0$ to obtain,

$$AsV_{CG}(\omega) = \mathbb{E}[X] - \sigma_X^2 \omega^2 + o(\omega^2) \quad (29)$$

which is the same as (10), expressed in terms of the mean and variance of X . When $\sigma_X^2 = 0$, X is a constant and η is Gaussian. If instead $\sigma_X^2 > 0$, then AsV can be improved by increasing ω in the neighborhood of 0, implying that $AsV_{pspc}^* < \sigma_\eta^2$.

As a concrete example, consider Middleton Class-A noise [17] where the variance RV is discrete and given by $X = \sigma_\eta^2 [Y/(A(T+1)) + T/(T+1)]$, A and T are deterministic parameters controlling the impulsiveness of the noise η , and Y is a Poisson RV with parameter A . In this case,

$$M_X(t) = \exp \left(t \frac{\sigma_\eta^2 T}{T+1} \right) \exp \left(A \left(\exp \left(\frac{t \sigma_\eta^2}{A(T+1)} \right) - 1 \right) \right). \quad (30)$$

Substituting in (28) we obtain the AsV . The resulting expression shows that when $T = 0$ (highly impulsive noise) $AsV_{CG}(\omega) \rightarrow 0$ as $\omega \rightarrow \infty$ in which case ω should be chosen as large as possible (i.e., $\omega = 2\pi/\theta_R$). Another interesting aspect of this expression is that it illustrates that $AsV(\omega)$ need not have a unique local minimum (i.e., it need not be convex or quasi-convex) for every sensing noise distribution. In fact, as will be seen in Figure 5 of the Simulations section, $AsV(\omega)$ can have multiple local minima, unlike the Gaussian, Cauchy and Laplace cases considered thus far.

C. Low Channel SNR Regime

When σ_v^2/P_T is sufficiently large, the $\varphi_\eta(2\omega)$ term in (7) is negligible, thereby transforming the problem in (8) into maximizing $[\omega\varphi_\eta(\omega)]^2$ over $(0, 2\pi/\theta_R]$. We now briefly summarize how the solutions in the previous subsection simplify in this regime. Since we already have closed form expressions for the solution of (8) for the Cauchy and Laplace cases, we only focus on the Gaussian and uniform cases.

For the Gaussian case maximizing $\omega^2 e^{-\sigma_\eta^2 \omega^2}$ over $\omega \in (0, 2\pi/\theta_R]$ yields $\omega^* = \min(2\pi/\theta_R, 1/\sigma_\eta)$. If θ_R is sufficiently small so that $\omega^* = 1/\sigma_\eta$, then we have

$$AsV_G(1/\sigma_\eta) = \frac{\sigma_\eta^2 e}{2} \left[\frac{\sigma_v^2}{P_T} + (1 - e^{-2}) \right] \quad (31)$$

which is an upper bound on the best achievable performance AsV^* , even when the channel SNR is not low, but becomes tighter at low channel SNR.

For the uniform case we maximize $\sin^2(\omega a)$ which yields $\omega^* = \min(2\pi/\theta_R, \pi/2a)$. If θ_R is small enough, $\omega^* = \pi/(2a)$ and $AsV_U(\pi/2a) = (a^2/2)(\sigma_v^2/P_T)$.

V. COMPARISON WITH AMPLIFY AND FORWARD SCHEME

In the AF scheme, the transmitted signal at the i^{th} sensor is $\alpha_L x_i$ where α_L depends on the number of sensors L to maintain the total power constraint, but is independent of x_i [10], [11]. We focus on the i.i.d. case for simplicity, and choose α_L identical across sensors due to symmetry. In what follows, we will show that the asymptotic performance of AF is competitive with that of the proposed scheme when the sensing noise has finite variance, and inferior to the proposed scheme when the sensing noise is impulsive.

The received signal for AF is,

$$y_L = \alpha_L \sum_{i=1}^L (\theta + \eta_i) + v. \quad (32)$$

We have already alluded to the fact that the per-sensor power $\alpha_L^2(\theta + \eta_i)^2$ is an unbounded RV, when the pdf of the sensing noise has infinite support. This is undesirable especially for low-power sensor networks with limited peak-power capabilities. Therefore, before we compare the asymptotic variances of the proposed estimator and AF, we reiterate that with respect to the management of the instantaneous transmit power of sensors, the proposed estimator is preferable to AF.

Since the total instantaneous power is random for AF, the total power is defined as an average $P_T = \alpha_L^2 \sum_{i=1}^L \mathbb{E}[(\theta + \eta_i)^2]$, with respect to the sensing noise distribution. We will consider a total power constraint case where P_T is not a function of L so that $\alpha_L = \sqrt{\frac{P_T}{L(\theta^2 + \sigma_\eta^2)}} = O(L^{-1/2})$.

The estimator in AF is given by $\hat{\theta}_{AF} = \text{Re}\{y_L\}/(L\alpha_L)$ so that

$$\sqrt{L}(\hat{\theta}_{AF} - \theta) = \frac{1}{\sqrt{L}} \sum_{i=1}^L \eta_i + \sqrt{\frac{\theta^2 + \sigma_\eta^2}{P_T}} \text{Re}\{v\} \quad (33)$$

with an AsV of,

$$AsV_{AF} = \sigma_\eta^2 + \frac{\sigma_v^2}{2P_T}(\theta^2 + \sigma_\eta^2) \quad (34)$$

when η has finite variance.

Consider now the special case of no channel noise ($\sigma_v^2 = 0$) which implies $AsV_{AF} = \sigma_\eta^2$. In Section IV-A we have seen that $AsV_{\text{pspc}}^* < \sigma_\eta^2$ is possible when the sensing noise is impulsive enough to have a positive excess kurtosis, the proposed approach outperforms AF when there is no channel noise. We now examine the more general case of $\sigma_v^2 > 0$.

Observe that (34) depends explicitly on θ , whereas (8) depends on the estimation range θ_R . Since it is difficult to compare these expressions in general, we will examine the case of large and small θ . When θ is large, $AsV_{AF} \approx \sigma_\eta^2 + (\sigma_v^2/P_T)(\theta^2/2)$, and by the discussion after (13), $AsV^* \approx \sigma_\eta^2 + (\sigma_v^2/P_T)(\theta_R^2/8\pi^2)$. Note that when the parameter θ is close to its upper limit, the proposed estimator will outperform AF. However, when θ is very small despite a large range θ_R , the AF will outperform the proposed approach.

Let us now examine the case of small θ and θ_R , where we focus on the Gaussian case. For this purpose, we bound the difference in performance between the proposed estimator and AF:

$$AsV^* - AsV_{AF} \leq AsV_G(1/\sigma_\eta) - AsV_{AF} \quad (35)$$

$$= \sigma_\eta^2 \left[\frac{e}{2}(1 - e^{-2}) - 1 + \frac{\sigma_v^2}{2P_T} \left(e - 1 - \frac{\theta^2}{\sigma_\eta^2} \right) \right], \quad (36)$$

where the inequality is because (31) is an upper bound on AsV^* . Examining the bound in (36) we note that its sign depends on the the channel SNR σ_v^2/P_T and the sensing SNR θ^2/σ_η^2 . In conclusion, the proposed approach is competitive with AF and may outperform it, depending on the specific parameter values when the sensing noise has finite variance. In what follows, the heavy-tailed sensing noise case is discussed.

With the AF approach the normalized multiple access channel output is proportional to the the sample mean, which is not a good estimator of θ when the sensing noise is heavy-tailed. To

illustrate with a specific example, consider the case when η is Cauchy. Dividing both sides of (33) with \sqrt{L} it is clear that $(\hat{\theta}_{AF} - \theta) \rightarrow 0$ is not possible since the sample mean $L^{-1} \sum_{i=0}^L \eta_i$ is Cauchy distributed and has the same distribution as η_1 regardless of the value of L . Since the sample mean is not a consistent estimator for Cauchy noise, the AF approach over multiple access channels fails for such a heavy-tailed distribution. On the other hand, the proposed estimator is strongly consistent in the presence of any noise distribution, including Cauchy. This brief example illustrates that the inherent robustness of our approach in the presence of heavy-tailed sensing noise distributions. The sample mean, “computed” by the multiple access channel in the AF approach, is highly suboptimal, and sometimes not consistent like in the Cauchy case, whereas in the proposed approach the channel computes (a noisy and normalized version of) the empirical characteristic function of the sensed samples, from which a consistent estimator can be constructed for any sensing noise distribution.

To be fair to AF, even though it suffers from having potentially large peak powers, we also want to point out the situations under which it is preferable to the proposed approach. The first point is that AF does not require the parameter θ to be bounded, and it does not require fine-tuning of a transmission parameter like ω . Moreover, AF is also a “universal” estimator, albeit over a smaller class of distributions (those that have finite variance) for the sensing noise.

In conclusion, the proposed estimator with its fixed instantaneous power per sensor is inherently preferable to AF when the sensors have a small dynamic range. Moreover, for AF, the total transmit power depends on θ and the statistics of the sensing noise. On the other hand, the AF approach has the benefit of not assuming θ to be in a finite set, and sometimes has a better finite sample performance as seen in the simulations. For impulsive noise distributions with finite variance and positive excess kurtosis like Laplace, or heavy-tailed distributions with infinite variance like Cauchy, the proposed approach is superior to AF. For other regimes, the two schemes are competitive and their asymptotic performance comparison depends on the specific values of parameters θ , θ_R , σ_v^2 , σ_η^2 , and P_T .

VI. FADING CHANNELS

Suppose that the multiple access channel connecting the sensors to the FC has fading so that (2) becomes

$$y_L = \sqrt{\rho} \sum_{i=1}^L |h_i| e^{j\omega x_i} + v \quad (37)$$

where $|h_i|$ is the amplitude of the channel coefficient h_i between the i^{th} sensor and the FC satisfying $E[|h_i|^2] = 1$. Even though the channel h_i is complex valued, the effective channel $|h_i|$ is real and positive when the i^{th} sensor corrects for the channel phase before transmission, using local channel phase information. Such a phase correction does not change the constant power nature of the transmission.

The following Theorem characterizes the performance of the proposed estimator over fading channels:

Theorem 8. *For the channel in (37) the estimator $\hat{\theta}$ in (6) is asymptotically normal with variance*

$$AsV(\omega) = (E[|h_i|])^{-2} \frac{\left[\frac{\sigma_v^2}{P_T} + 1 - \varphi_\eta(2\omega) \right]}{2\omega^2 \varphi_\eta^2(\omega)} \quad (38)$$

Proof: The proof is similar to that of Theorem 1 with the following changes:

$$\begin{aligned} v_c &= \frac{1}{2} + \frac{1}{2} \varphi_\eta(\omega) - (E[|h_i|])^2 \varphi_\eta^2(\omega) \\ v_s &= \frac{1}{2} - \frac{1}{2} \varphi_\eta(2\omega), \end{aligned}$$

and both G_1 and G_2 are scaled by a factor of $(E[|h_i|])^{-1}$. Substituting these in (42) we obtain (38). ■

Since $E[|h_i|^2] = 1$, using Jensen's inequality, the $(E[|h_i|])^{-2}$ factor due to fading is always less than one, unless $|h_i|$ is deterministic. In fact, when $|h_i|$ is Ricean the loss due to fading is given by $(\sqrt{K+1} \Gamma(3/2) e^{-K} {}_1F_1(3/2; 1; K))^{-2}$ where ${}_1F_1(\cdot; \cdot; \cdot)$ is the confluent hypergeometric function [18, pp. 504] and K is the Ricean parameter. This expression reduces to $4/\pi$ when $K = 0$, implying Rayleigh fading channels. In the AF setting, the difference between fading and no fading also exhibits the same loss, which was analyzed in detail in [4], [5] for different fading distributions, where the Nakagami case was also considered. Note that if the optimization of the asymptotic variance is desired in the fading case, the fading loss does not affect the optimum value of ω so equations (17), (20), (23), and (26) remain valid for the different sensing noise distributions.

VII. SIMULATIONS

In what follows, we corroborate our analytical results through Monte Carlo simulations, and also examine finite-sample effects that are not predictable from our asymptotic results.

In Figures 1 and 2 we compare $AsV(\omega)$ and $L\text{var}(\hat{\theta} - \theta)$ versus ω for the per-sensor, and total power constraints, respectively. We begin by acknowledging that the variance of the asymptotic distribution, $AsV(\omega)$, and the normalized limiting variance $L\text{var}(\hat{\theta} - \theta)$ are not always equal in general [19, pp. 437]. However, as the next two figures show, they are in agreement for the proposed estimator. The mismatch that occurs for small ω are due to the number of samples not being sufficiently large for both Figures 1 and 2. To focus more on this mismatch, in Figures 3 and 4 we consider smaller values of L , and an increased range for ω for the Gaussian sensing noise case. As expected, for reduced values of L the mismatch increases, especially for small, and large values of ω . Note that for the per-sensor power constraint case, although $AsV(\omega)$ is bounded near the origin, with finite samples, $L\text{var}(\hat{\theta} - \theta)$ is large for small ω , an effect which is more pronounced for small L . This suggests that for the per-sensor power constraint case ω should not be chosen arbitrarily small, especially when L is small, to avoid this finite-sample artifact.

In Figure 5 we compare $AsV(\omega)$ and $L\text{var}(\hat{\theta} - \theta)$ versus ω for the per-sensor, and total power constraints, respectively, for Middleton Class A noise. In addition to the agreement of the theory and simulations, these plots illustrate that $AsV(\omega)$ need not be a convex, or a quasiconvex function of ω with a unique local minimum. For all the other noise distributions, $AsV(\omega)$ did exhibit a unique local minimum, which was helpful in finding the optimal value of ω .

Figures 6 and 7 show $L\text{var}(\hat{\theta} - \theta)$ versus L for the per-sensor, and total power constraints, respectively. The optimal value of ω that minimizes the AsV is chosen for the total power constraint case. For the per-sensor power constraint case in Figure 6, we did not use the minimizer of $AsV(\omega)$ due to the aforementioned finite-sample effects. Instead, the value of ω is chosen to minimize $L\text{var}(\hat{\theta} - \theta)$ in Figure 1 (which assumes $L = 500$) and applied to all values of L in Figure 6. It is seen that convergence occurs slower for the heavy-tailed Cauchy distribution. At about $L = 50$, all cases converge for both Figures. Figure 8 illustrates the effect of Rayleigh fading on the performance for Gaussian sensing and channel noise. It is seen that $L\text{var}(\hat{\theta} - \theta)$ converges to their theoretically predicted asymptotic value with a ratio of about $4/\pi$ compared

to the non-fading case.

In Figure 9 the proposed scheme is compared with AF. The performance of the proposed approach is seen to be both better and worse than AF depending on the value of θ . Another interesting aspect of Figure 9 is the flatness of the curves for the AF case. This can be seen by finding the variance of equation (33), which is a constant function of L . In contrast, the normalized variance for the proposed estimator is seen to depend on L in Figure 9.

To illustrate the robustness of the proposed estimator Figure 10 compares it with AF for Cauchy sensing noise. One realization of the estimation error is plotted for each value of L to illustrate that in the presence of Cauchy noise, the performance of AF does not converge despite the increase in L , whereas the proposed estimator is consistent.

VIII. CONCLUSIONS

A distributed estimation scheme relying on constant modulus transmissions from the sensors is proposed over Gaussian multiple access channels. The instantaneous transmit power does not depend on the random sensing noise, which is a desirable feature for low-power sensors with limited peak power capabilities. In the i.i.d. case, the estimator is shown to be strongly consistent for any sensing or channel noise distribution. In the non-identically distributed case, a bound on the variances is shown to be a sufficient condition for strong consistency. The asymptotic variance is derived, and shown to depend on the characteristic function of the sensing noise which is bounded for the general case, and also optimized with respect to ω for various noise distributions. In addition to the desirable constant-power feature, the proposed estimator is robust to impulsive noise, and remains consistent even when the mean and variance of the sensing noise does not exist. It is argued that over Gaussian multiple access channels, the AF estimator is effectively a noisy sample mean of the sensed data. For sensing noise distributions for which the sample mean is highly suboptimal or inconsistent, the proposed estimator is shown to outperform AF. The effect of fading is also considered, and shown to effect the asymptotic variance by a constant fading penalty factor.

APPENDIX 1: PROOF OF THEOREM 4

We begin by observing that the 2×1 vector sequence, $\sqrt{L} [z_L^R - \bar{z}^R \quad z_L^I - \bar{z}^I]$ is asymptotically normal with zero mean, due to the central limit theorem. The elements of its asymptotic

covariance matrix can be calculated to be,

$$\Sigma_{11} := P_T [\cos^2(\omega\theta) \cdot v_c + \sin^2(\omega\theta) \cdot v_s] + \frac{\sigma_v^2}{2} \quad (39)$$

$$\Sigma_{22} := P_T [\cos^2(\omega\theta) \cdot v_s + \sin^2(\omega\theta) \cdot v_c] + \frac{\sigma_v^2}{2} \quad (40)$$

$$\Sigma_{12} = \Sigma_{21} := P_T \sin(\omega\theta) \cos(\omega\theta) (v_c - v_s) \quad (41)$$

where, for brevity we have $v_c := \text{var}(\cos \omega \eta_i) = (1/2) + \varphi_\eta(2\omega)/2 - \varphi_\eta^2(\omega)$ and $v_s := \text{var}(\sin \omega \eta_i) = (1/2) - \varphi_\eta(2\omega)/2$. Applying [14, Thm 3.16] the asymptotic variance is given by

$$AsV = G_1^2 \Sigma_{11} + 2G_1 G_2 \Sigma_{12} + G_2^2 \Sigma_{22} \quad (42)$$

where

$$G_1 := \left. \frac{\partial \frac{1}{\omega} \tan^{-1} \frac{x}{y}}{\partial x} \right|_{x=\bar{z}^R} = -\frac{1}{\omega} \frac{1}{1 + \tan^2(\omega\theta)} \frac{\tan(\omega\theta)}{\sqrt{P_T} \varphi_\eta(\omega) \cos(\omega\theta)} \quad (43)$$

$$G_2 := \left. \frac{\partial \frac{1}{\omega} \tan^{-1} \left(\frac{x}{y} \right)}{\partial y} \right|_{y=\bar{z}^I} = \frac{-G_1}{\tan \omega\theta} \quad (44)$$

Substituting in (42) and simplifying we obtain the theorem.

APPENDIX 2: PROOF OF THEOREM 7

Using the bound, $\varphi_{\eta_i}(\omega) \geq 1 - \sigma_{\eta_i}^2 \omega^2 / 2$, we have for all ω , $\varphi_\eta(2\omega) \geq 1 - 2\sigma_\eta^2 \omega^2$, and for $\omega < \sqrt{2}/\sigma_\eta$, $\varphi_\eta^2(\omega) \geq (1 - \sigma_\eta^2 \omega^2 / 2)^2$. Substituting in (7) we have for $\omega < \sqrt{2}/\sigma_\eta$,

$$AsV(\omega) \leq \frac{\left[\frac{\sigma_\eta^2}{P_T} + 2\sigma_\eta^2 \omega^2 \right]}{2\omega^2 \left(1 - \frac{\sigma_\eta^2 \omega^2}{2} \right)^2}. \quad (45)$$

Recall that $2\pi/\theta_R < \sqrt{2}/\sigma_\eta$ by assumption. Therefore, upper bound (45) is valid over the entire range of ω values which involves the minimization in (8). We can therefore minimize both sides of (45) over $\omega \in (0, 2\pi/\theta_R]$. Substituting for convenience $\beta \leftarrow \omega^2 \sigma_\eta^2$ we have

$$AsV^* \leq \min_{\beta \in (0, (2\pi\sigma_\eta/\theta_R)^2]} \frac{\left[\frac{\sigma_v^2}{P_T} + 2\beta \right]}{2\beta(1 - \beta/2)^2} \sigma_\eta^2. \quad (46)$$

The unconstrained minimum can be found by differentiating (46) and is given by $c/8$, with a corresponding minimum given by the right hand side (rhs) of (12). It can be checked that $c/8$ is the unique minimum of the unconstrained problem. This shows that if $(2\pi\sigma_\eta/\theta_R)^2 > c/8$ then the rhs of (46) is given by the rhs of (12).

To show (13), recall that $c/8$ is the unique unconstrained minimum of the objective on the rhs of (46). This implies that as a function of β it is non-increasing over $(0, c/8)$ so that when $(2\pi\sigma_\eta/\theta_R)^2 < c/8$ the minimum over $[0, (2\pi\sigma_\eta/\theta_R)^2]$ is achieved at $\beta = (2\pi\sigma_\eta/\theta_R)^2$ which is the rhs of (13).

REFERENCES

- [1] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 27–41, July 2006.
- [2] M. Gastpar, "Uncoded transmission is exactly optimal for a simple gaussian sensor network," *IEEE Transaction on Information Theory*, vol. 54, pp. 5247–5251, Nov 2008.
- [3] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 757–770, Feb. 2008.
- [4] M. K. Banavar, C. Tepedelenlioglu, and A. Spanias, "Performance of distributed estimation over multiple access fading channels with partial feedback," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, Las Vegas, NV, Mar./Apr. 2008, pp. 2253–2256.
- [5] C. Tepedelenlioglu, M. K. Banavar, and A. Spanias, "Asymptotic analysis of distributed estimation over fading multiple access channels," in *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*, Pacific Grove, CA, Nov. 2007, pp. 2140–2144.
- [6] T. Wimalajeewa and S. K. Jayaweera, "Power efficient distributed estimation in a bandlimited wireless sensor network," in *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*, Pacific Grove, CA, Nov. 2007, pp. 2156–2160.
- [7] G. Mergen and L. Tong, "Type based estimation over multiaccess channels," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 613–626, Feb. 2006.
- [8] S. Marano, V. Matta, L. Tong, and P. Willett, "Bandwidth scaling for efficient inference over a power-limited MAC," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3, Honolulu, HI, Apr. 2007, pp. 597–600.
- [9] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2210–2219, June 2005.
- [10] S. Cui, J. J. Xiao, A. J. Goldsmith, Z. Q. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4683–4695, 2007.
- [11] M. Gastpar and M. Vetterli, "Source-Channel communication in sensor networks." International Workshop on Information Processing in Sensor Networks (IPSN'03), March 2003, pp. 162–177.
- [12] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol I*. John Wiley & Sons, 1968.
- [13] N. G. Ushakov, *Selected topics in characteristic functions*. VSP International Science Publishers, 1999.
- [14] B. Porat, *Digital processing of random signals: theory and methods*. Prentice-Hall, Englewood Cliffs, NJ, 1994.
- [15] R. Zamir, "A proof of the Fisher information inequality via a data processing argument," *Information Theory, IEEE Transactions on*, vol. 44, no. 3, pp. 1246–1250, May 1998.

- [16] T. C. Scott and R. B. Mann, “General relativity and quantum mechanics: Towards a generalization of the Lambert W Function,” *Journal reference: AAECC (Applicable Algebra in Engineering, Communication and Computing)*, vol. 16, no. 6, 2006.
- [17] S. Zabin and H. Poor, “Parameter estimation for Middleton Class A interference processes,” *IEEE Transactions on Communications*, vol. 37, no. 10, pp. 1042–1051, Oct 1989.
- [18] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. Courier Dover Publications, 1965.
- [19] E. L. Lehmann, G. Casella, and G. Casella, *Theory of point estimation*. Springer New York, 1998.

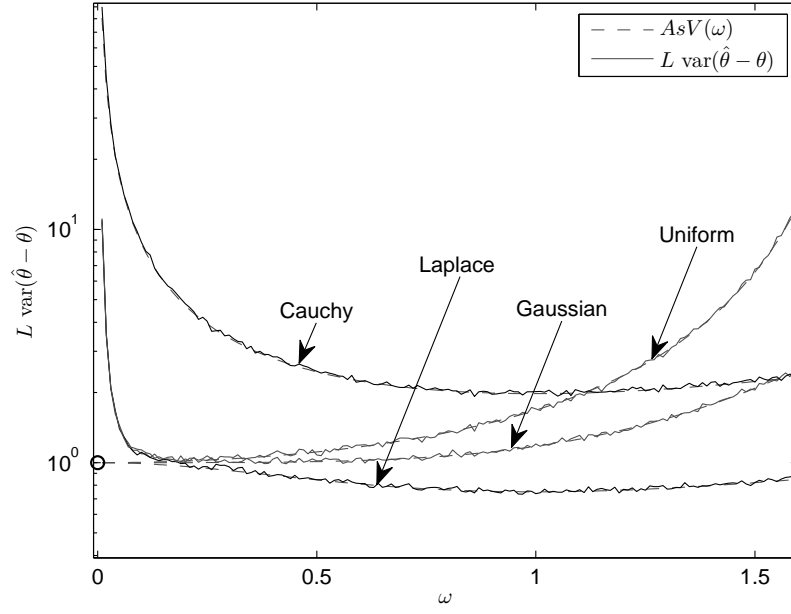


Fig. 1. Per-sensor Power Constraint, $\sigma_\eta^2 = 1$, $\sigma_v^2 = 1$, $L = 500$, $\rho = 1$

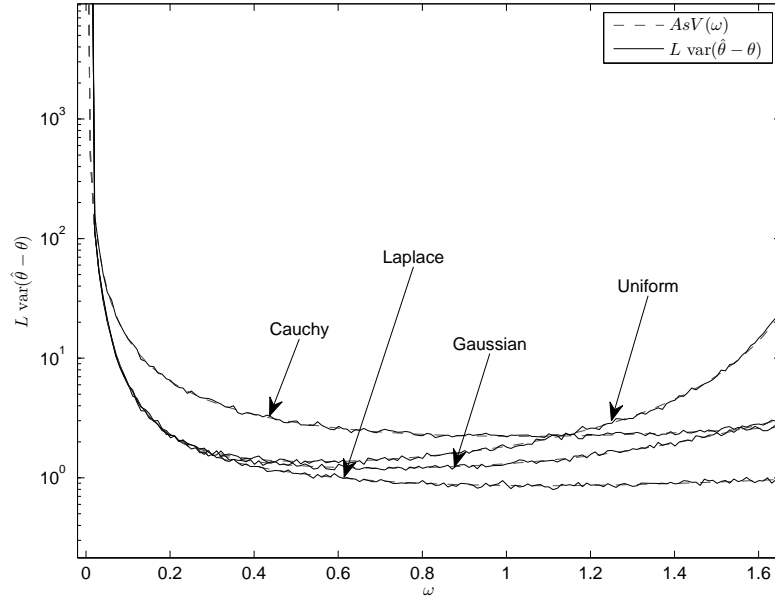


Fig. 2. Total Power Constraint, $\sigma_\eta^2 = 1$, $\sigma_v^2 = 1$, $L = 500$, $P_T = 10$

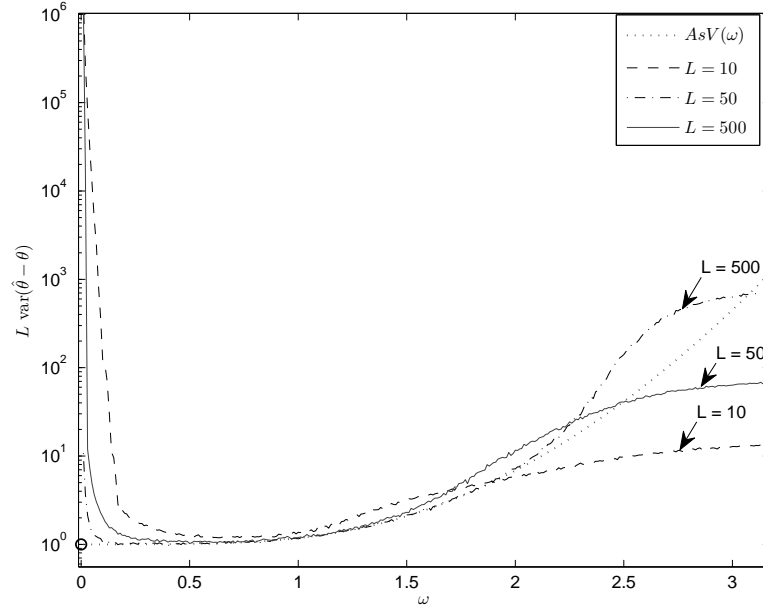


Fig. 3. Per-sensor Power Constraint, $\sigma_\eta^2 = 1$, $\sigma_v^2 = 1$, $\rho = 1$

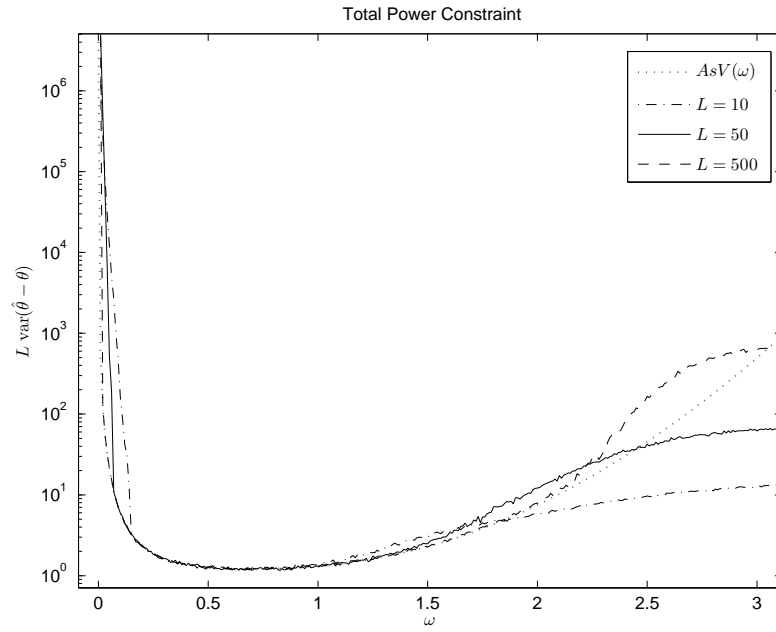


Fig. 4. Total Power Constraint, $\sigma_\eta^2 = 1$, $\sigma_v^2 = 1$, $P_T = 10$

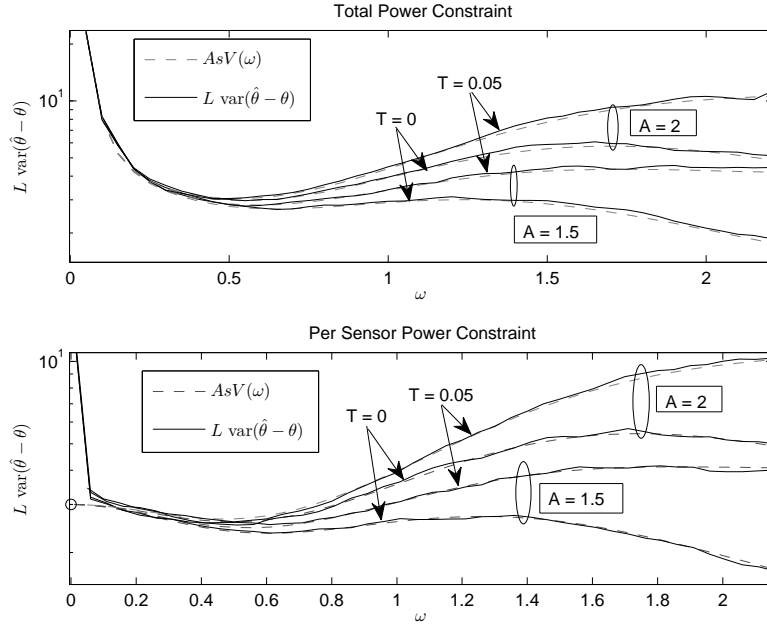


Fig. 5. $\sigma_\eta^2 = 3$, $\sigma_v^2 = 1$, $L = 500$, $P_T = 10$

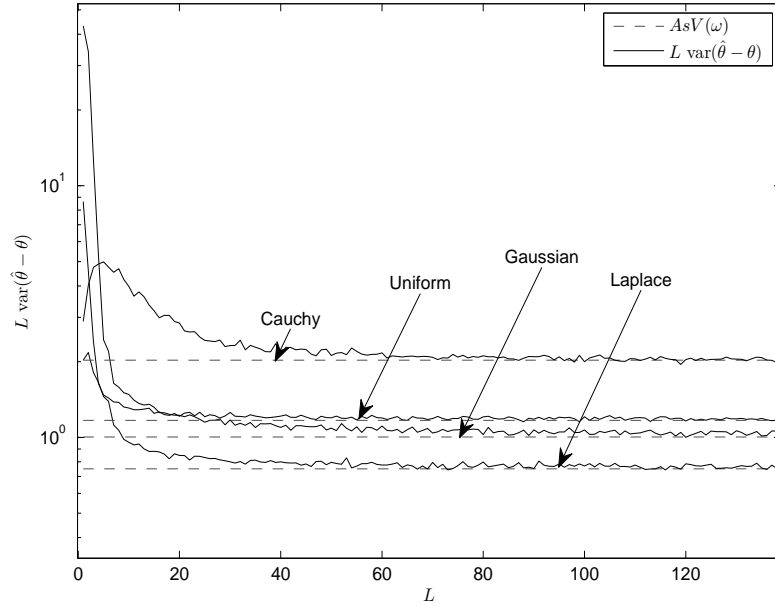


Fig. 6. Per-sensor Power Constraint, $\sigma_\eta^2 = 1$, $\sigma_v^2 = 1$, $\rho = 1$, $\theta_R = 4$, $\theta = 2$

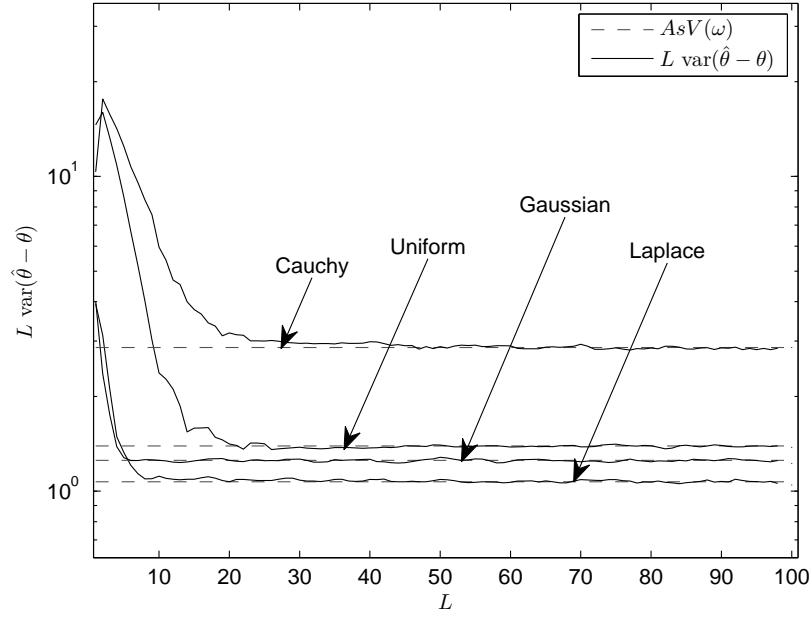


Fig. 7. Total Power Constraint, $\sigma_\eta^2 = 1$, $\sigma_v^2 = 1$, $P_T = 10$, $\theta_R = 12$, $\theta = 2$

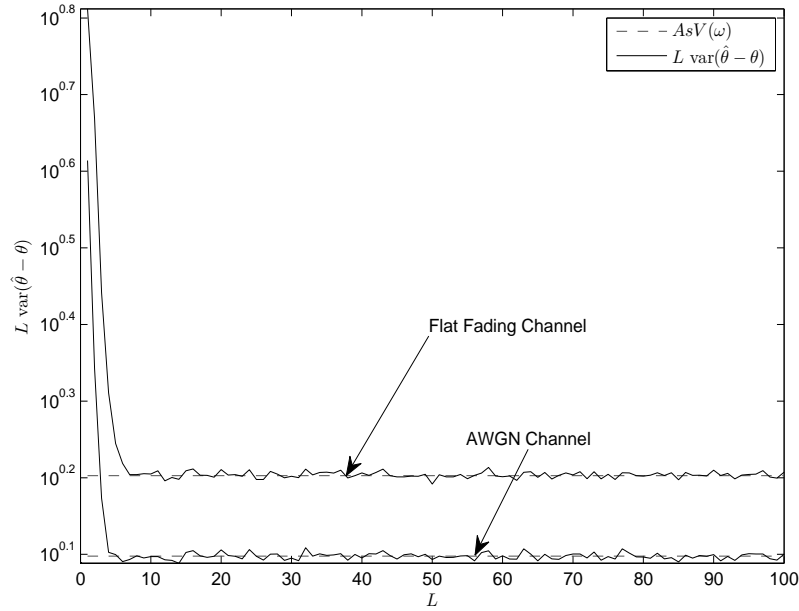


Fig. 8. Total Power Constraint, $E[|h|^2] = 1$, $\sigma_\eta^2 = 1$, $\sigma_v^2 = 1$, $P_T = 10$, $\theta_R = 12$, $\theta = 2$,

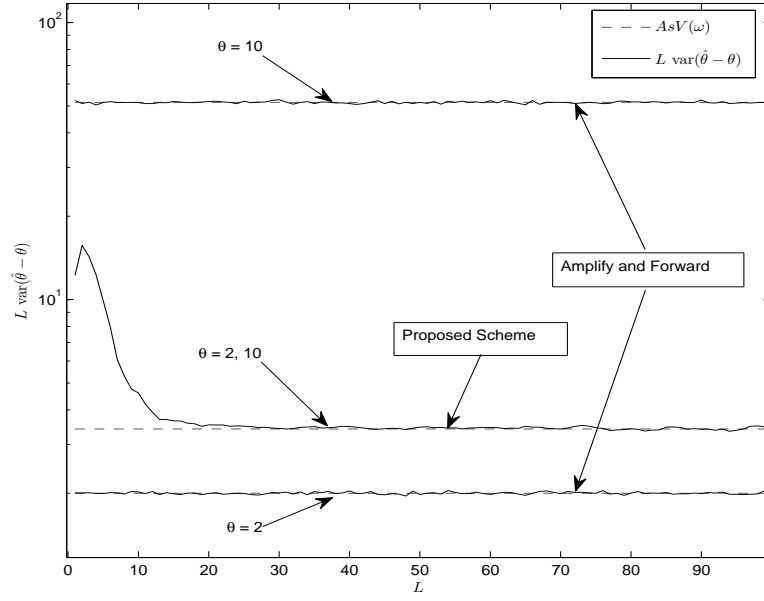


Fig. 9. Total Power Constraint, $\sigma_\eta^2 = 1$, $\sigma_v^2 = 1$, $P_T = 10$, $\theta_R = 12$

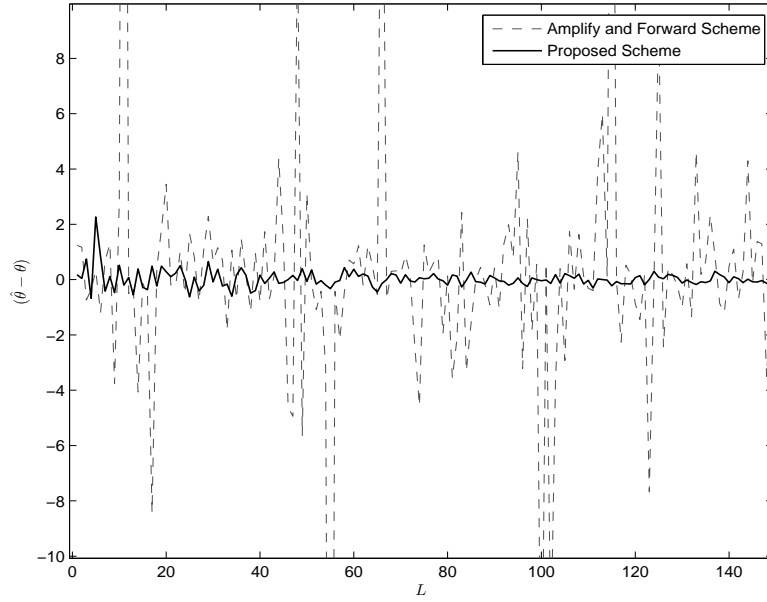
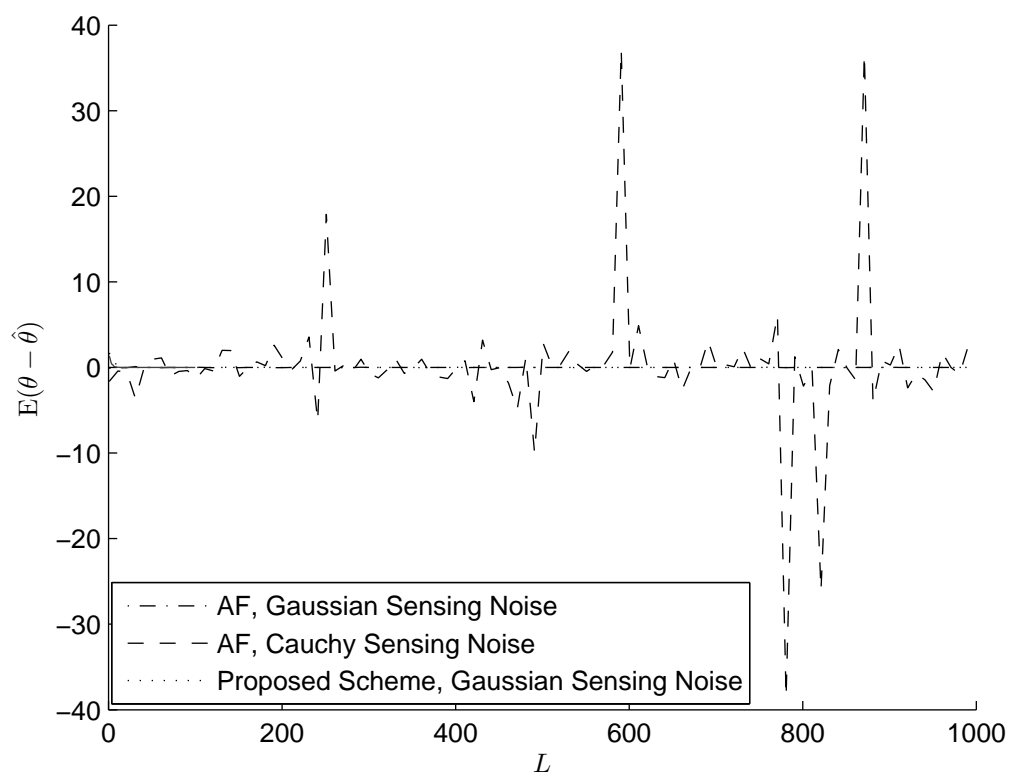
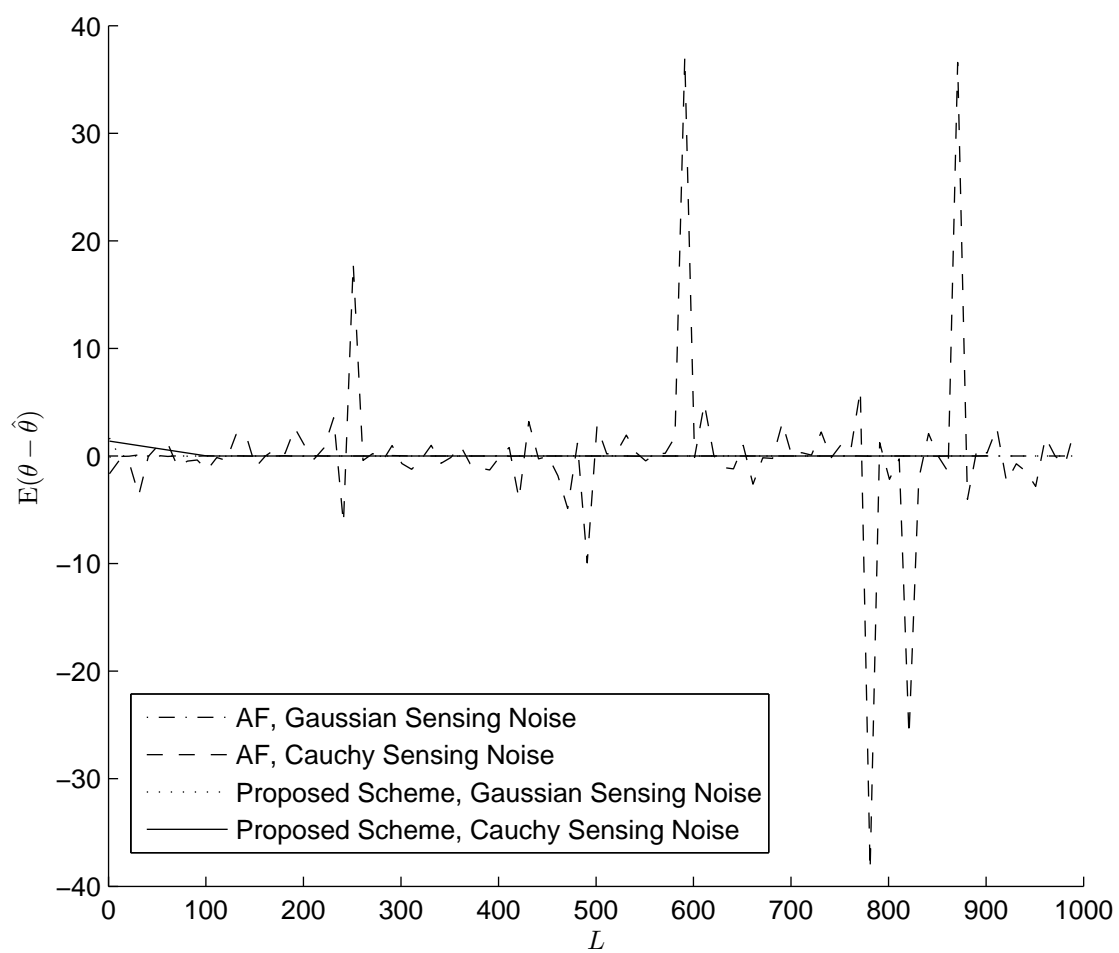
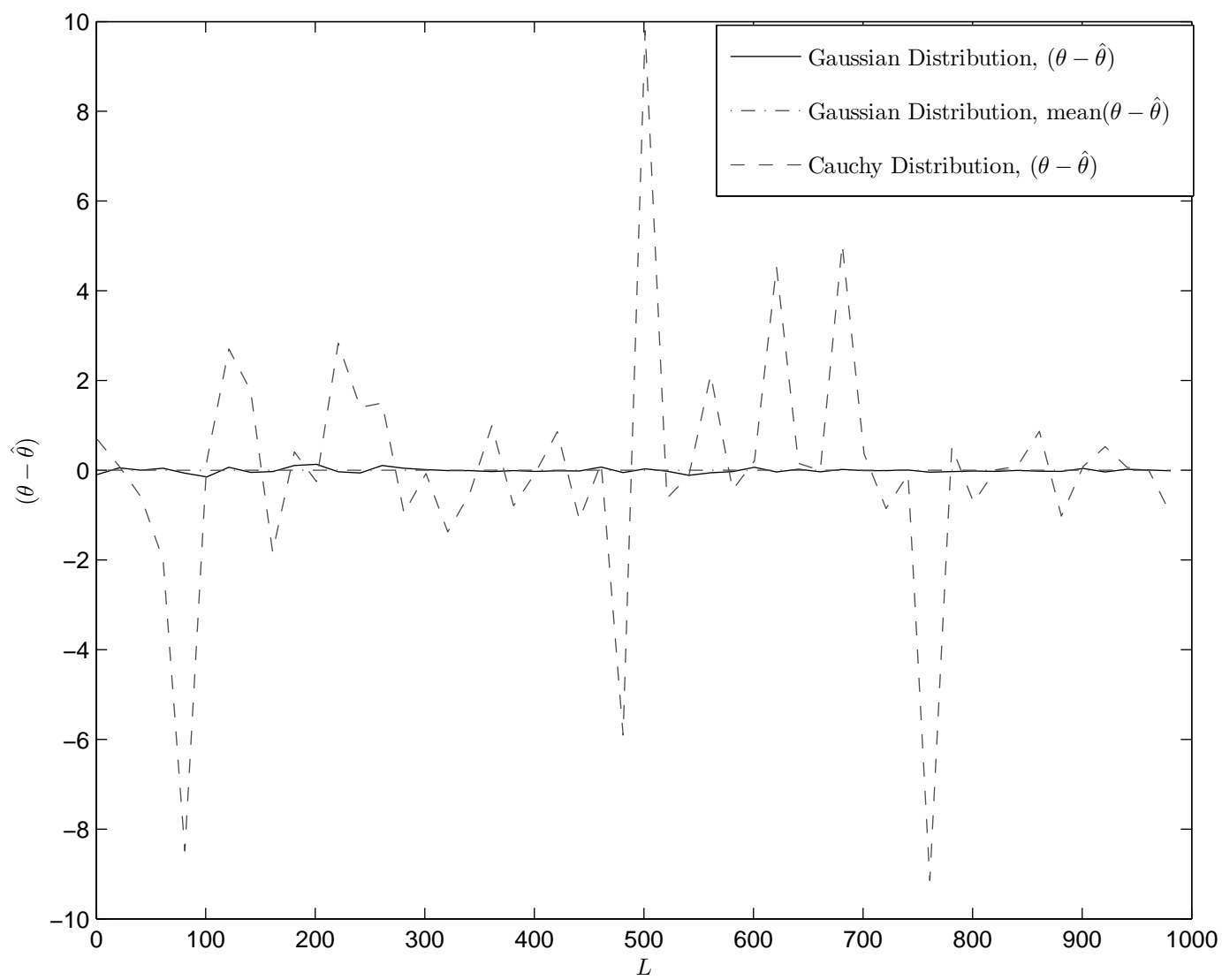
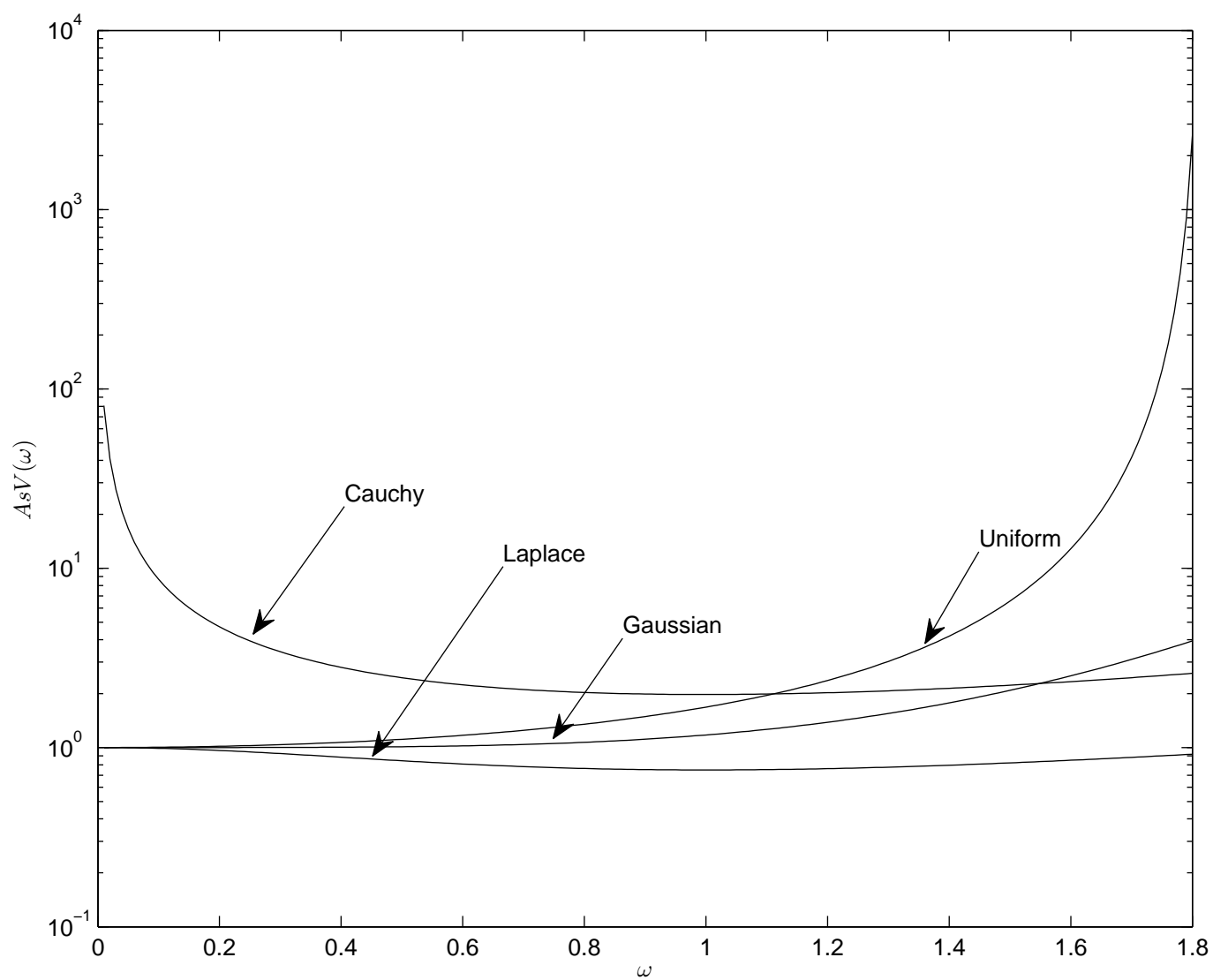


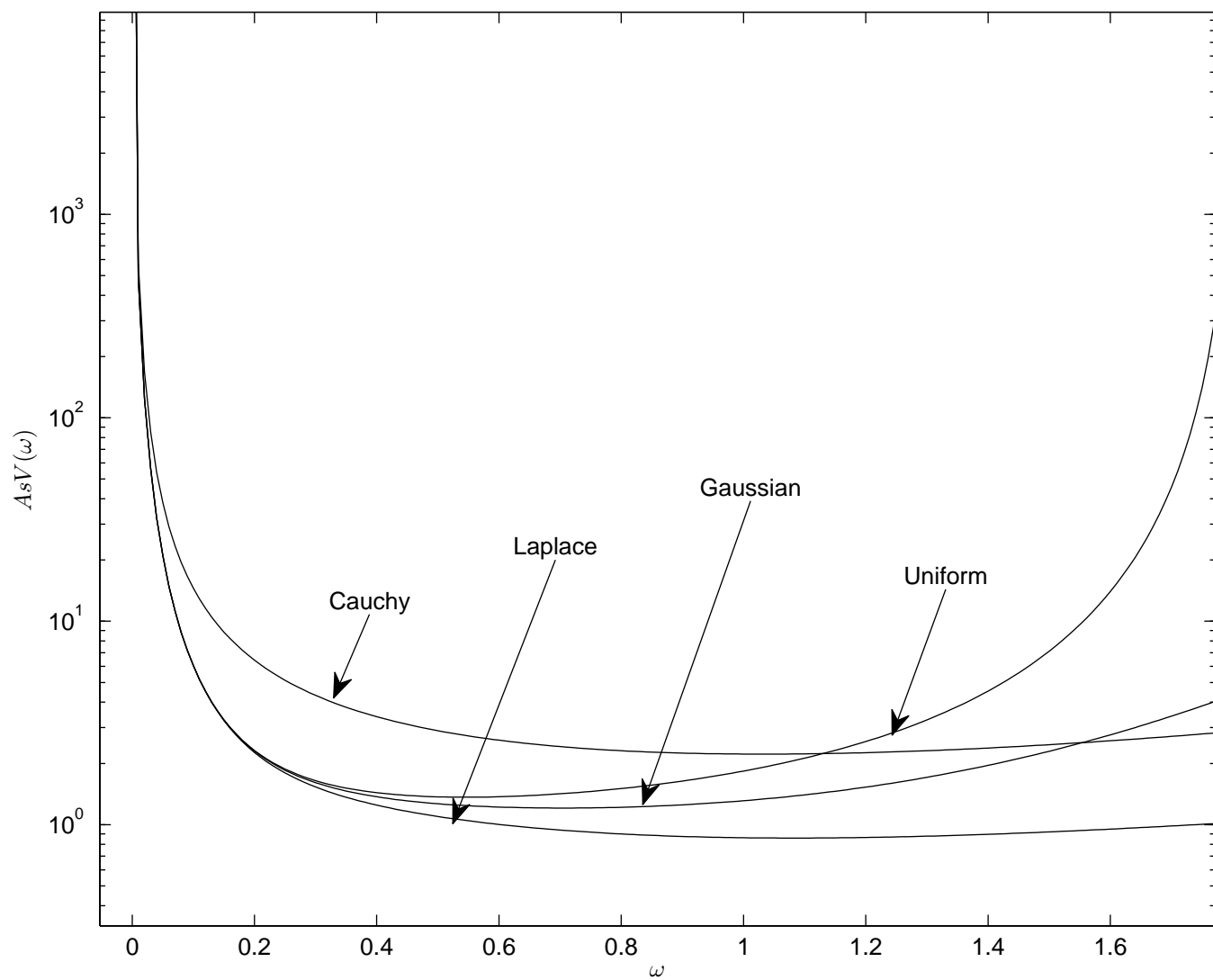
Fig. 10. Total Power Constraint with Cauchy distributed sensing noise, $\sigma_\eta^2 = 1$, $\sigma_v^2 = 1$, $P_T = 10$, $\theta_R = 12$, $\theta = 2$



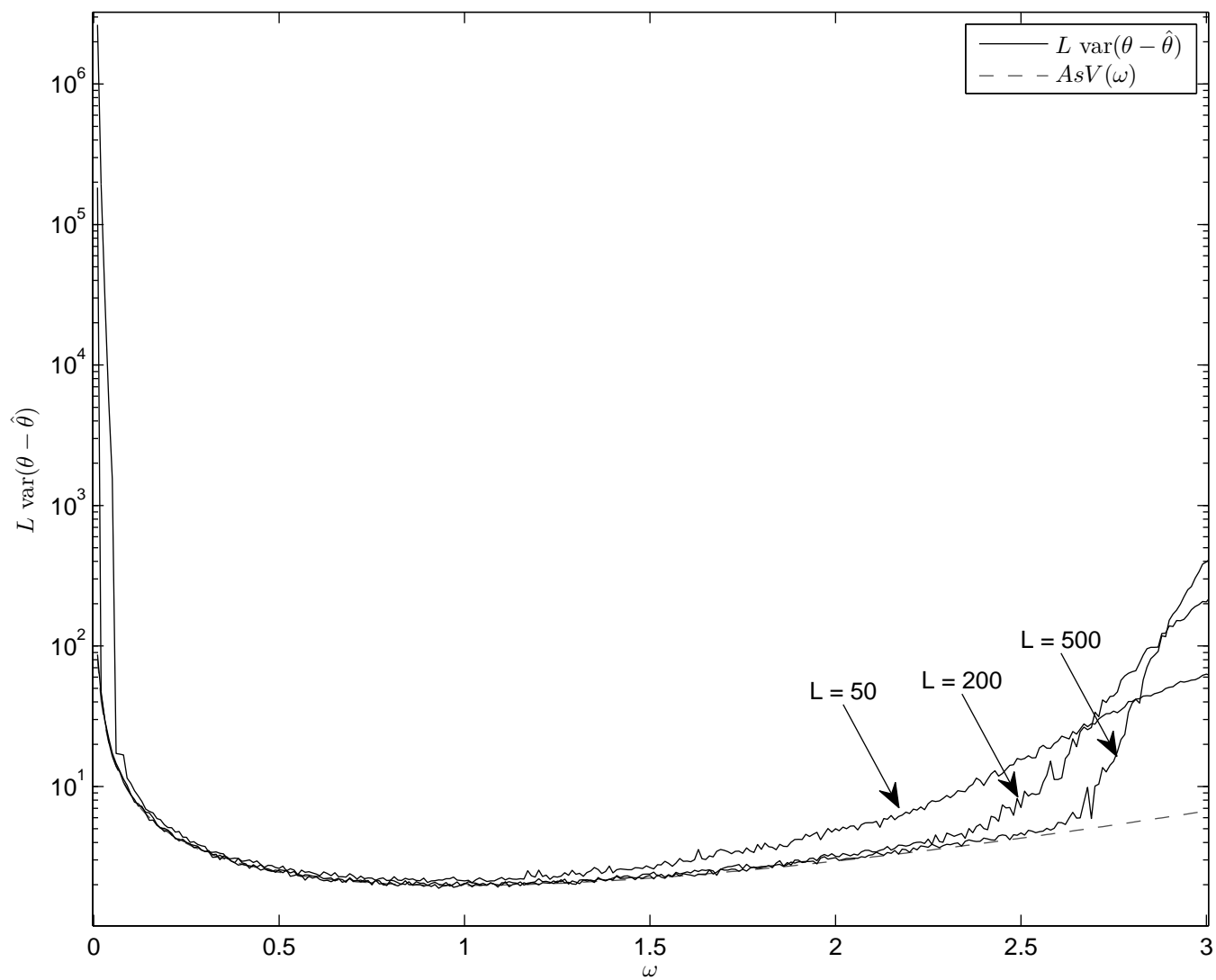


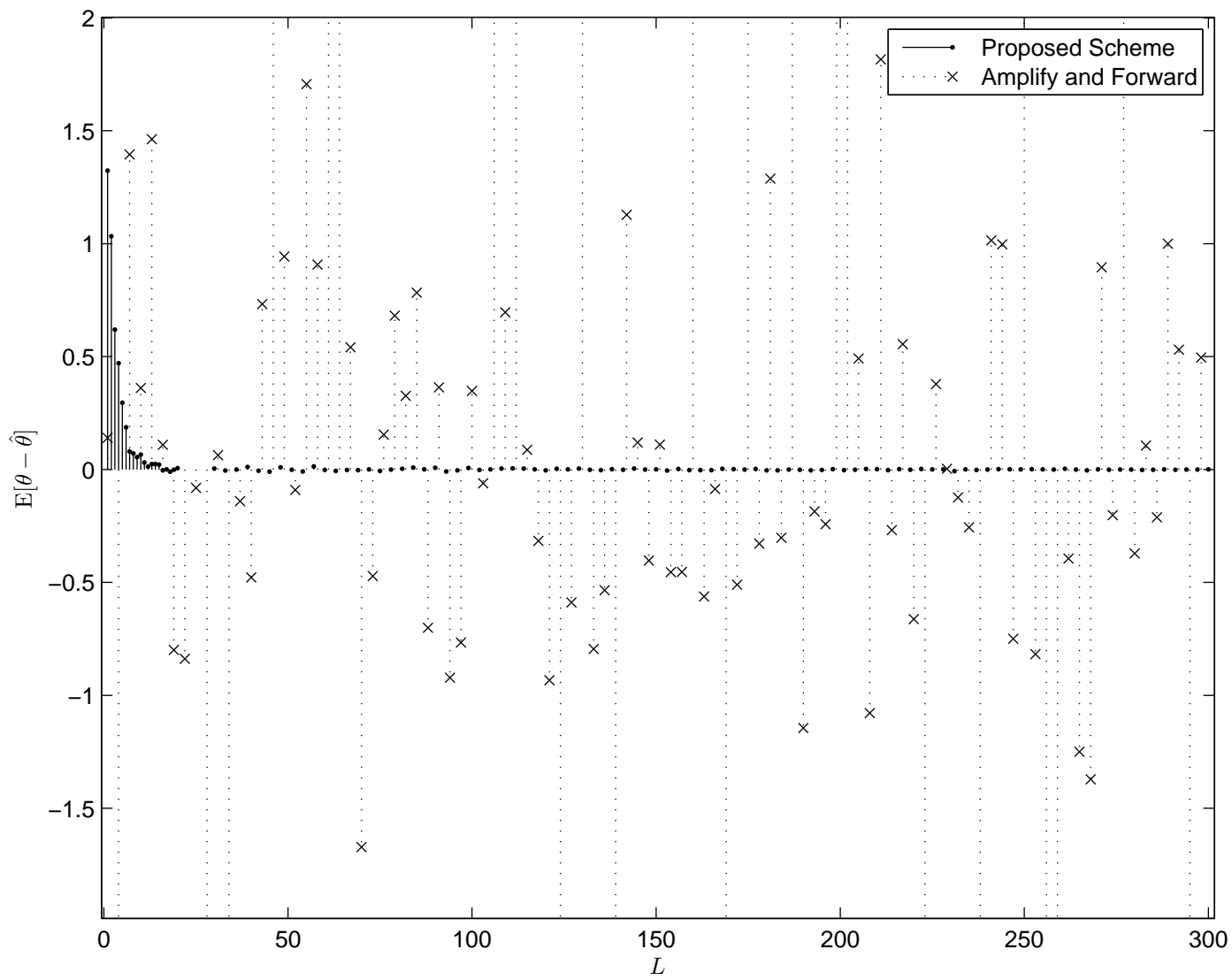




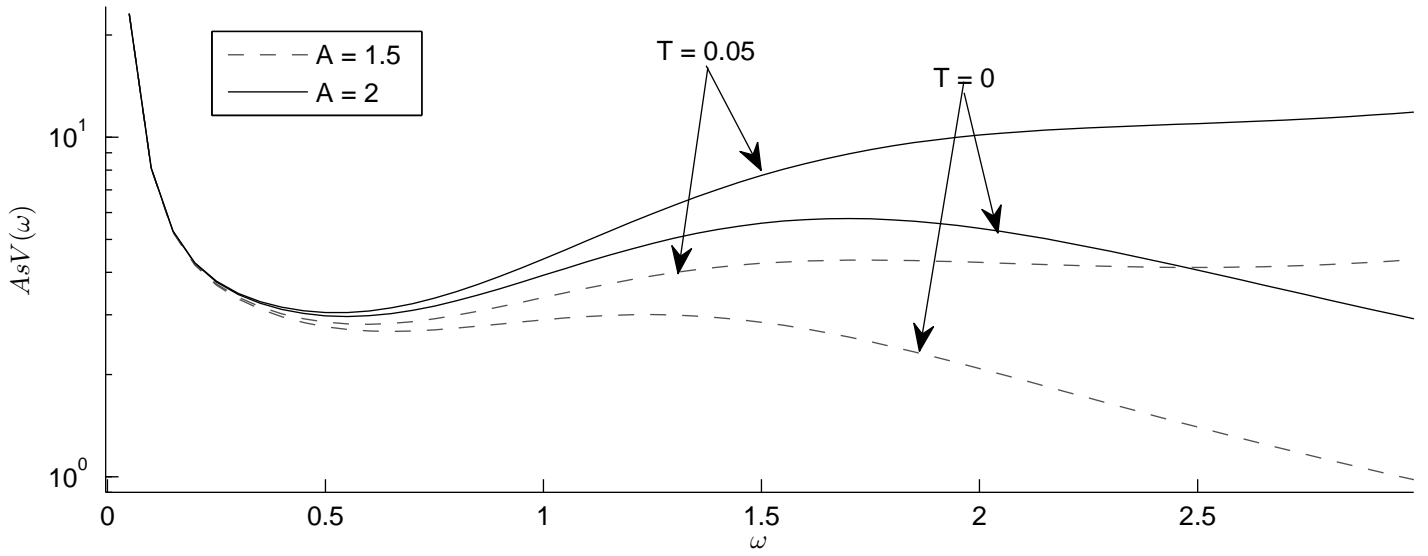


Cauchy Distributed Sensing Noise





Total Power Constraint



Per Sensor Power Constraint

